PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/00 | A2

(11) International Publication Number: **WO 98/37094**

(43) International Publication Date: 27 August 1998 (27.08.98)

(21) International Application Number: PCT/US98/03595

(22) International Filing Date: 24 February 1998 (24.02.98)

(30) Priority Data:

08/804,561	24 February 1997 (24.02.97)	US
09/028,168	23 February 1998 (23.02.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

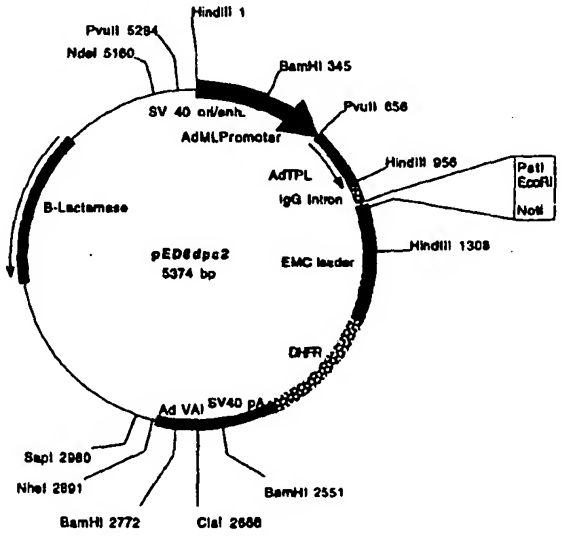
Published

Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

10

This application is a continuation-in-part of Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application 08/804,561), filed February 24, 1997, which is incorporated by reference herein.

15

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

20

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly
25 related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning
30 techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 65 to nucleotide 1270;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1139 to nucleotide 1270;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1011 to nucleotide 1216;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BO114_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BO114_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BO114_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BO114_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 196 to amino acid 205 of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 65 to nucleotide 1270; the nucleotide sequence of SEQ ID NO:1 from nucleotide 1139 to nucleotide 1270; the nucleotide sequence of SEQ ID NO:1 from nucleotide 1011 to nucleotide 1216; the nucleotide sequence of the full-length protein coding sequence of clone BO114_1 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone BO114_1 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BO114_1 deposited under accession number ATCC 98333. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 326 to amino acid 384.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 326 to amino acid 384;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 196 to amino acid 205 of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone BO114_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 326 to amino acid 384.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 418 to nucleotide 582;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 508 to nucleotide 582;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 555;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CD311_2 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CD311_2 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CD311_2 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CD311_2 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 22 to amino acid 31 of SEQ ID NO:4;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 418 to nucleotide 582; the nucleotide sequence of SEQ ID NO:3 from nucleotide 508 to nucleotide 582; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 555; the nucleotide sequence of the full-length protein coding sequence of clone CD311_2 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone CD311_2 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of

clone CD311_2 deposited under accession number ATCC 98333. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 46.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 46;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 22 to amino acid 31 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone CD311_2 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 46.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 191 to nucleotide 1756;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 254 to nucleotide 1756;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 604;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CG279_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CG279_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CG279_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CG279_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 256 to amino acid 265 of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 191 to nucleotide 1756; the nucleotide sequence of SEQ ID NO:5 from nucleotide 254 to nucleotide 1756; the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 604; the nucleotide sequence of the full-length protein coding sequence of clone CG279_1 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone CG279_1 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CG279_1 deposited under accession number ATCC 98333. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 138.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 138;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 256 to amino acid 265 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone CG279_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 138.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 226 to nucleotide 948;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1128 to nucleotide 1601;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CJ424_9 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CJ424_9 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CJ424_9 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CJ424_9 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 115 to amino acid 124 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 226 to nucleotide 948; the nucleotide sequence of SEQ ID NO:7 from nucleotide 1128 to nucleotide 1601; the nucleotide sequence of the full-length protein coding sequence of clone CJ424_9 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone CJ424_9 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CJ424_9 deposited under accession number ATCC 98333.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 115 to amino acid 124 of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone CJ424_9 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 137 to nucleotide 895;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1488 to nucleotide 2274;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CR930_1 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CR930_1 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CR930_1 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CR930_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 121 to amino acid 130 of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 137 to nucleotide 895; the nucleotide sequence of SEQ ID NO:9 from nucleotide 1488 to nucleotide 2274; the nucleotide sequence of the full-length protein coding sequence of clone CR930_1 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone CR930_1 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CR930_1 deposited under accession number ATCC 98333.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 121 to amino acid 130 of SEQ ID NO:10; and

(c) the amino acid sequence encoded by the cDNA insert of clone CR930_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 494 to nucleotide 973;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 611 to nucleotide 973;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 521 to nucleotide 940;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DA306_4 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DA306_4 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DA306_4 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DA306_4 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 75 to amino acid 84 of SEQ ID NO:12;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 494 to nucleotide 973; the nucleotide sequence of SEQ ID NO:11 from nucleotide 611 to nucleotide 973; the nucleotide sequence of SEQ ID NO:11 from nucleotide 521 to nucleotide 940; the nucleotide sequence of the full-length protein coding sequence of clone DA306_4 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone DA306_4 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DA306_4 deposited under accession number ATCC 98333. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 11 to amino acid 149.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 11 to amino acid 149;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 75 to amino acid 84 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone DA306_4 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 11 to amino acid 149.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 2295 to nucleotide 2594;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1867 to nucleotide 2372;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DG76_1 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DG76_1 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DG76_1 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DG76_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 45 to amino acid 54 of SEQ ID NO:14;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 2295 to nucleotide 2594; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1867 to nucleotide 2372; the nucleotide sequence of the full-length protein coding sequence of clone DG76_1 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone DG76_1 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DG76_1 deposited under accession number ATCC 98333. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 26.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 26;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 45 to amino acid 54 of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone DG76_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 26.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 394 to nucleotide 522;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 476;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DO19_1 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DO19_1 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DO19_1 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DO19_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 394 to nucleotide 522; the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 476; the nucleotide sequence of the full-length protein coding sequence of clone DO19_1 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone DO19_1 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone DO19_1 deposited under accession number ATCC 98333. In yet other preferred

embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 27.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 27;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone DO19_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 27.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 262 to nucleotide 654;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 322 to nucleotide 654;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 618;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone EQ219_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone EQ219_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone EQ219_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone EQ219_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 60 to amino acid 69 of SEQ ID NO:18;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 262 to nucleotide 654; the nucleotide sequence of SEQ ID NO:17 from nucleotide 322 to nucleotide 654; the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 618; the nucleotide sequence of the full-length protein coding sequence of clone EQ219_1 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone EQ219_1 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone EQ219_1 deposited under accession number ATCC 98333. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 119.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 119;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 60 to amino acid 69 of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone EQ219_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18 or the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 119.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 74 to nucleotide 310;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 125 to nucleotide 310;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 1 to nucleotide 338;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone FG340_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone FG340_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone FG340_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone FG340_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 34 to amino acid 43 of SEQ ID NO:20;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 74 to nucleotide 310; the nucleotide sequence of SEQ ID NO:19 from nucleotide 125 to nucleotide 310; the nucleotide sequence of SEQ ID NO:19 from nucleotide 1 to nucleotide 338; the nucleotide sequence of the full-length protein coding sequence of clone FG340_1 deposited under accession number ATCC 98333; or the nucleotide sequence of the mature protein coding sequence of clone FG340_1 deposited under accession number ATCC 98333. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone FG340_1 deposited under accession number ATCC 98333. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 75.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 75;

(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising the amino acid sequence from amino acid 34 to amino acid 43 of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone FG340_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 75.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide

sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "BO114_1"

A polynucleotide of the present invention has been identified as clone "BO114_1". BO114_1 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BO114_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BO114_1 protein").

The nucleotide sequence of BO114_1 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BO114_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 346 to 358 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 359, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BO114_1 should be approximately 1600 bp.

The nucleotide sequence disclosed herein for BO114_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BO114_1 demonstrated at least some similarity with sequences

identified as AA430329 (zw20e04.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 769854 5' similar to contains element MER22 repetitive element), H00825 (yj31b05.r1 Homo sapiens cDNA clone 150321 5'), H12557 (yj12c09.r1 Homo sapiens cDNA clone 148528 5'), W53899 (md09c01.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA), and Z82202 (Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 34P24). Based upon sequence similarity, BO114_1 proteins and each similar protein or peptide may share at least some activity.

Clone "CD311_2"

A polynucleotide of the present invention has been identified as clone "CD311_2". CD311_2 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CD311_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CD311_2 protein").

The nucleotide sequence of CD311_2 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CD311_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 18 to 30 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 31, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CD311_2 should be approximately 2400 bp.

The nucleotide sequence disclosed herein for CD311_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CD311_2 demonstrated at least some similarity with sequences identified as H17421 (ym40d12.s1 Homo sapiens cDNA clone 50810 3'), H20618 (ym47b01.r1 Homo sapiens cDNA clone 51411 5'), and U70476 (Rattus norvegicus cationic amino acid transporter-1 (CAT-1) mRNA, complete cds). Based upon sequence similarity, CD311_2 proteins and each similar protein or peptide may share at least some activity.

Clone "CG279_1"

A polynucleotide of the present invention has been identified as clone "CG279_1". CG279_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CG279_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CG279_1 protein").

The nucleotide sequence of CG279_1 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CG279_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 9 to 21 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 22, or are a transmembrane domain. Amino acids 43 to 55 are a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 56, or are a transmembrane domain

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CG279_1 should be approximately 3940 bp.

The nucleotide sequence disclosed herein for CG279_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CG279_1 demonstrated at least some similarity with sequences identified as AA568111 (nf13c05.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:913640), D63222 (Human placenta cDNA 5'-end GEN-508F12), H64777 (yu62h09.r1 Homo sapiens cDNA clone 238433 5' similar to contains Alu repetitive element;contains TAR1 repetitive element), M17262 (Human prothrombin (F2) gene, complete cds, and Alu and KpnI repeats), Q39724 (Expressed Sequence Tag human gene marker EST00316), U14685 (Gorilla gorilla Alu-Sb2 repeat, clone GO-14), U14691 (Gorilla gorilla Alu-Sb2 repeat, clone GOI2-11), and W15458 (zc19h02.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 322803 3'). The predicted amino acid sequence disclosed herein for CG279_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CG279_1 protein demonstrated at least some similarity to sequences identified as D25215 (KIAA0032 [Homo sapiens]), M15530 (B-cell growth factor [Homo sapiens]), and R95913 (Neural thread protein). Based upon sequence similarity, CG279_1 proteins and each similar

protein or peptide may share at least some activity. The TopPredII computer program predicts three potential transmembrane domains within the CG279_1 protein sequence, centered around amino acids 30, 250, and 390 of SEQ ID NO:6, respectively. The nucleotide sequence of CG279_1 indicates that it may contain one or more Alu repetitive elements.

Clone "CJ424_9"

A polynucleotide of the present invention has been identified as clone "CJ424_9". CJ424_9 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CJ424_9 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CJ424_9 protein").

The nucleotide sequence of CJ424_9 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CJ424_9 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CJ424_9 should be approximately 1650 bp.

The nucleotide sequence disclosed herein for CJ424_9 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CJ424_9 demonstrated at least some similarity with sequences identified as AB000215 (Rattus norvegicus cca1 mRNA, complete cds) and R83763 (yp16f06.s1 Homo sapiens cDNA clone 187619 3'). The predicted amino acid sequence disclosed herein for CJ424_9 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CJ424_9 protein demonstrated at least some similarity to sequences identified as AB000215 (CCA1 protein [Rattus norvegicus]) and M59465 (A20 [Homo sapiens]). Based upon sequence similarity, CJ424_9 proteins and each similar protein or peptide may share at least some activity.

Clone "CR930_1"

A polynucleotide of the present invention has been identified as clone "CR930_1". CR930_1 was isolated from a human adult testes cDNA library using methods which are

selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CR930_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CR930_1 protein").

The nucleotide sequence of CR930_1 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CR930_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CR930_1 should be approximately 2400 bp.

The nucleotide sequence disclosed herein for CR930_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CR930_1 demonstrated at least some similarity with sequences identified as AA058338 (zk82e08.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 489350 3'), N54229 yz03f05.r1 Homo sapiens cDNA clone), R89733 (ym99e10.r1 Homo sapiens cDNA clone 167082 5'), and W60141 (zc94f06.s1 Pancreatic Islet Homo sapiens cDNA clone 338819 3'). The predicted amino acid sequence disclosed herein for CR930_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CR930_1 protein demonstrated at least some similarity to sequences identified as a *C. elegans* ORF (open reading frame) (U21324) that is weakly similar to *S. cerevisiae* CBP3 protein precursor (SP:CBP3_YEAST, P21560), a mitochondrial membrane protein. Based upon sequence similarity, CR930_1 proteins and each similar protein or peptide may share at least some activity.

Clone "DA306_4"

A polynucleotide of the present invention has been identified as clone "DA306_4". DA306_4 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DA306_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "DA306_4 protein").

The nucleotide sequence of DA306_4 as presently determined is reported in SEQ ID NO:11. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DA306_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12. Amino acids 27 to 39 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 40, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DA306_4 should be approximately 2800 bp.

The nucleotide sequence disclosed herein for DA306_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DA306_4 demonstrated at least some similarity with sequences identified as AA397398 (nc65a07.r1 NCI CGAP Pr1 Homo sapiens cDNA clone 771444), AL008629 (Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 197O17; HTGS phase 1), D78769 (Human placenta cDNA 5'-end GEN-512A03), M19364 (Human gamma-B-crystallin (gamma 1-2) and gamma-C-crystallin (gamma 2-1) genes, complete cds), N29380 (yw97f09.s1 Homo sapiens cDNA clone 260201 3'), N47928 (yw97f09.r1 Homo sapiens cDNA clone 260201 5'), Z83745 (Human DNA sequence from PAC 453A3 contains EST and STS), and Z83848 (Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 57A13; HTGS phase 1). The predicted amino acid sequence disclosed herein for DA306_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted DA306_4 protein demonstrated at least some similarity to sequences identified as M12140 (envelope protein [Homo sapiens]), M19051 (pol protein [Mus musculus]), R75189 (Osteoinductive retrovirus RFB-14 pol gene product), U88902 (integrase [Homo sapiens]). Based upon sequence similarity, DA306_4 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of DA306_4 indicates that it may contain a CpG island repeat region.

Clone "DG76_1"

A polynucleotide of the present invention has been identified as clone "DG76_1". DG76_1 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DG76_1 is a full-length clone,

including the entire coding sequence of a secreted protein (also referred to herein as "DG76_1 protein").

The nucleotide sequence of DG76_1 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DG76_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DG76_1 should be approximately 3300 bp.

The nucleotide sequence disclosed herein for DG76_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DG76_1 demonstrated at least some similarity with sequences identified as AA044352 (zk54c01.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 486624 5') and N57171 (yw90f09.r1 Homo sapiens cDNA clone 259529 5'). Based upon sequence similarity, DG76_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the DG76_1 protein sequence, centered amino acids 15 and 80 of SEQ ID NO:14, respectively. The nucleotide sequence of DG76_1 indicates that it may contain a MER repeat region.

Clone "DO19_1"

A polynucleotide of the present invention has been identified as clone "DO19_1". DO19_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. DO19_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "DO19_1 protein").

The nucleotide sequence of DO19_1 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the DO19_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone DO19_1 should be approximately 700 bp.

The nucleotide sequence disclosed herein for DO19_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. DO19_1 demonstrated at least some similarity with sequences identified as AA339440 (EST44546 Fetal brain I Homo sapiens cDNA 5' end). Based upon sequence similarity, DO19_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of DO19_1 indicates that it may contain one or more MER20 repetitive elements.

Clone "EQ219_1"

A polynucleotide of the present invention has been identified as clone "EQ219_1". EQ219_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. EQ219_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "EQ219_1 protein").

The nucleotide sequence of EQ219_1 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the EQ219_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Amino acids 8 to 20 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 21, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone EQ219_1 should be approximately 800 bp.

The nucleotide sequence disclosed herein for EQ219_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. EQ219_1 demonstrated at least some similarity with sequences identified as AA400429 (zu62a09.s1 Soares testis NHT Homo sapiens cDNA clone 742552 3'). Based upon sequence similarity, EQ219_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts an additional potential transmembrane domain within the EQ219_1 protein sequence, centered around amino acid 90 of SEQ ID NO:18.

Clone "FG340_1"

A polynucleotide of the present invention has been identified as clone "FG340_1". FG340_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. FG340_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "FG340_1 protein").

The nucleotide sequence of FG340_1 as presently determined is reported in SEQ ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the FG340_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20. Amino acids 5 to 17 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 18, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone FG340_1 should be approximately 900 bp.

The nucleotide sequence disclosed herein for FG340_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. FG340_1 demonstrated at least some similarity with sequences identified as N59424 (yv51g05.s1 Homo sapiens cDNA clone 246296 3'). The FG340_1 nucleotide sequence has an interesting simple "TG" nucleotide repeat from basepair 96 to basepair 131 of SEQ ID NO:19. This region encodes a Cys-Val repeat in the FG340_1 protein. Similar Cys-Val stretches are found in the amino termini of X52164 (Q300 protein (AA 1-77) [Mus musculus]) and M37679 (Ig heavy chain precursor [Mus musculus]). Based upon sequence similarity, FG340_1 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones BO114_1, CD311_2, CG279_1, CJ424_9, CR930_1, DA306_4, DG76_1, DO19_1, EQ219_1, and FG340_1 were deposited on February 20, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98333, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the

deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* **9**: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
BO114_1	SEQ ID NO:21
CD311_2	SEQ ID NO:22
CG279_1	SEQ ID NO:23
CJ424_9	SEQ ID NO:24
CR930_1	SEQ ID NO:25
DA306_4	SEQ ID NO:26
DG76_1	SEQ ID NO:27

DO19_1	SEQ ID NO:28
EQ219_1	SEQ ID NO:29
FG340_1	SEQ ID NO:30

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination,

preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologs are those isolated from mammalian species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides .

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1551 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTGTTATTGT AGCGAGGGAC ATGAGCTGGA GGCTGATGGC ATCAGCTGCA GCCCTGCAGG 60

GGCCATGGGT GCCCAGGCTT CCCAGGACCT CGGAGATGAG TTGCTGGATG ACGGGGAGGA 120

TGAGGAAGAT GAAGACGAGG CCTGGAAGGC CTTCAACGGT GGCTGGACGG AGATGCCTGG 180

GATCCTGTGG ATGGAGCCTA CGCAGCCGCC TGACTTTGCC CTGGCCTATA GACCGAGCTT 240

CCCAGAGGAC AGAGAGCCAC AGATACCCTA CCCGGAGCCC ACCTGGCCAC CCCCGCTCAG 300

TGCCCCCAGG GTCCCCTACC ACTCCTCAGT GCTCTCCGTC ACCCGGCCTG TGGTGGTCTC 360

TGCCACGCGT CCCACACTGC CTTCTGCCCA CCAGCCTCCT GTGATCCCTG CCACACACCC 420

AGCTTTGTCC CGTGACCACC AGATCCCCGT GATCGCAGCC AACTATCCAG ATCTGCCTTC 480

TGCCTACCAA CCCGGTATTC TCTCTGTCTC TCATTCAGCA CAGCCTCCTG CCCACCAGCC 540

CCCTATGATC TCAACCAAAT ATCCGGAGCT CTTCCCTGCC CACCAGTCCC CCATGTTTCC 600

AGACACCCGG GTCGCTGGCA CCCAGACCAC CACTCATTTG CCTGGAATCC CACCTAACCA 660

TGCCCCTCTG GTCACCACCC TCGGTGCCCA GCTACCCCCT CAAGCCCCAG ATGCCCTTGT 720

CCTCAGAACC CAGGCCACCC AGCTTCCCAT TATCCCAACT GCCCAGCCCT CTCTGACCAC 780

CACCTCCAGG TCCCCTGTGT CTCCTGCCCA TCAAATCTCT GTGCCTGCTG CCACCCAGCC 840

CGCAGCCCTC CCCACCCTCC TGCCCTCTCA GAGCCCCACT AACCAGACCT CACCCATCAG 900

CCCTACACAT CCCCATTCCA AAGCCCCCCA AATCCCAAGG GAAGATGGCC CCAGTCCCAA 960

GTTGGCCCTG TGGCTGCCCT CACCAGCTCC CACAGCAGCC CCAACAGCCC TGGGGGAGGC 1020

TGGTCTTGCC GAGCACAGCC AGAGGGATGA CCGGTGGCTG CTGGTGGCAC TCCTGGTGCC 1080

AACGTGTGTC TTTTTGGTGG TCCTGCTTGC ACTGGGCATC GTGTACTGCA CCCGCTGTGG 1140

CCCCCATGCA CCCAACAAGC GCATCACTGA CTGCTATCGC TGGGTCATCC ATGCTGGGAG 1200

CAAGAGCCCA ACAGAACCCA TGCCCCCCAG GGGCAGCCTC ACAGGGGTGC AGACCTGCAG 1260

AACCAGCGTG TGATGGGGTG CAGACCCCCC TCATGGAGTA TGGGGCGCTG GACACATGGC 1320

CGGGGCTGCA CCAGGGACCC ATGGGGGCTG CCCAGCTGGA CAGATGGCTT CCTGCTCCCC 1380

AGGCCCAGCC AGGGTCCTCT CTCAACCACT AGACTTGGCT CTCAGGAACT CTGCTTCCTG 1440

GCCCAGCGCT CGTGACCAAG GATACACCAA AGCCCTTAAG ACCTCAGGGG GCGGGTGCTG 1500

GGGTCTTCTC CAATAAATGG GGTGTCAACC TTAAAAAAAA AAAAAAAAAA A 1551

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 402 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Ala Gln Ala Ser Gln Asp Leu Gly Asp Glu Leu Leu Asp Asp
1 5 10 15

Gly Glu Asp Glu Glu Asp Glu Asp Glu Ala Trp Lys Ala Phe Asn Gly
20 25 30

Gly Trp Thr Glu Met Pro Gly Ile Leu Trp Met Glu Pro Thr Gln Pro
35 40 45

Pro Asp Phe Ala Leu Ala Tyr Arg Pro Ser Phe Pro Glu Asp Arg Glu
50 55 60

Pro Gln Ile Pro Tyr Pro Glu Pro Thr Trp Pro Pro Pro Leu Ser Ala
65 70 75 80

Pro Arg Val Pro Tyr His Ser Ser Val Leu Ser Val Thr Arg Pro Val
85 90 95

Val Val Ser Ala Thr Arg Pro Thr Leu Pro Ser Ala His Gln Pro Pro
100 105 110

Val Ile Pro Ala Thr His Pro Ala Leu Ser Arg Asp His Gln Ile Pro
115 120 125

Val Ile Ala Ala Asn Tyr Pro Asp Leu Pro Ser Ala Tyr Gln Pro Gly
130 135 140

Ile Leu Ser Val Ser His Ser Ala Gln Pro Pro Ala His Gln Pro Pro
145 150 155 160

Met Ile Ser Thr Lys Tyr Pro Glu Leu Phe Pro Ala His Gln Ser Pro
165 170 175

Met Phe Pro Asp Thr Arg Val Ala Gly Thr Gln Thr Thr Thr His Leu
180 185 190

```
Pro Gly Ile Pro Pro Asn His Ala Pro Leu Val Thr Thr Leu Gly Ala
        195                 200                 205

Gln Leu Pro Pro Gln Ala Pro Asp Ala Leu Val Leu Arg Thr Gln Ala
    210                 215                 220

Thr Gln Leu Pro Ile Ile Pro Thr Ala Gln Pro Ser Leu Thr Thr Thr
225                 230                 235                 240

Ser Arg Ser Pro Val Ser Pro Ala His Gln Ile Ser Val Pro Ala Ala
                245                 250                 255

Thr Gln Pro Ala Ala Leu Pro Thr Leu Leu Pro Ser Gln Ser Pro Thr
            260                 265                 270

Asn Gln Thr Ser Pro Ile Ser Pro Thr His Pro His Ser Lys Ala Pro
        275                 280                 285

Gln Ile Pro Arg Glu Asp Gly Pro Ser Pro Lys Leu Ala Leu Trp Leu
    290                 295                 300

Pro Ser Pro Ala Pro Thr Ala Ala Pro Thr Ala Leu Gly Glu Ala Gly
305                 310                 315                 320

Leu Ala Glu His Ser Gln Arg Asp Asp Arg Trp Leu Leu Val Ala Leu
                325                 330                 335

Leu Val Pro Thr Cys Val Phe Leu Val Val Leu Leu Ala Leu Gly Ile
            340                 345                 350

Val Tyr Cys Thr Arg Cys Gly Pro His Ala Pro Asn Lys Arg Ile Thr
        355                 360                 365

Asp Cys Tyr Arg Trp Val Ile His Ala Gly Ser Lys Ser Pro Thr Glu
    370                 375                 380

Pro Met Pro Pro Arg Gly Ser Leu Thr Gly Val Gln Thr Cys Arg Thr
385                 390                 395                 400

Ser Val
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2473 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGAGTATTGT GTCTACTTTT ATCTGTGCAC CAGCCACAAA TACCCACATT GGAAAGACCC 60

ATTTGTGATG GGTAAACATC CCTTCCTGTC TCCCACAACC CCTGTGACTG CCCTGCATGT 120

GTTCATGACC TCCGAAGGCC CAAATTCATG AAGCAGCAAA CCCAGCAGAT CTCCACCCCC 180

CTGCCTCAGG ACCTCTGCTG AAGAGGGGGA TGAAGTGGGT CTCCAGGGAG GCAGTGGGGG 240

CCTTGTTGGC AGCTGGCTCG GGAGCCGGCT TACAGGAGGG CAGCTCTGCA GTTGGGAGGG 300

GCACCGTCCG GAGGAGACCA GGCCTCTACA CACCCCCCAC TCTACTTATC ATCCCTGCTC 360

ACACACCCTT GTCCAAGGCT TTATGCATCG GATTTATTTT TCCAAATCAA GAGGACAGTG 420

ATAGATGCAT TTTCCCCAGG CTGTCTCAGA AAGGTCGCTA AATGTATACT GTTGTCAGAA 480

TTGCTGAGAT CTCCCCCCAC TTTTGGTTTT TGCAGCAGTA AAAACTCTTT CCACTGTGAC 540

TTATTTTCTC TCTCAGGCAG CCAGCCACCT GGTCCCTTGT GCTGACTCTA GCACAGTGGC 600

CAGGATCCAA TACGAGTCCA GGGGTGACCG CAGGATGGTG GGGGCAGCGG GCTTCTCCAC 660

CTACCCCAGC CACCAAGGCC CTGACGCACT GCCTCCTGCA CCTTCAGCAC ATCCCTGTGC 720

ACAGCTGGAA GGGTGCATGG CCCGCTCACC TTTGTTCAGA TGGGTGGAAA CGCTGATGAT 780

ACCAGCTCCT CCCTGCCGTG CCCCTGCCAC GGAGCAGGCA TTGTGAACTG GCTGGTGTTT 840

GCAGTCCCAC GTGGCATGGC CTCCAGCCCA ACCCACAGTG GAGACTGGAG ACAGGGCAAT 900

GAGTCTGGTC GGGGGCACGT GGACATGCCC CATAGGGGCC CCACCCAGAC TTAACAGGCA 960

AGGTCCTGGG CATTGCGCGA CGCAGGACTC AATGCTAAAG CAAGCCTGCC TGGCTCTGTG 1020

CCAGGGCCCC TCTTCTGATT TACACATCCC ATTTTTACAC AGACCCTTCC TTCTTAATAA 1080

AGGCTGACAG TTCTGTTGGC AGCCAAGAAC CCACACCATG AAGACAGGGA GTGAGGGGCC 1140

TTTGTGCCCA ACTCCAGCAC AGCTGCGTTC TGGGGTGTGT GAGAGGCATG TTCGTGTCTG 1200

TGCGCTGGTG GTCTCGTGAG ACAGTTCCGA GGACGGGGAA ATTGCAGGGT GGTGGGGGCG 1260

TGAGGCTTAT ATGTGGAACT GATGCAGAGT TCGCCTGCAG ACGGATCTGG ATATACACTA 1320

TGTATAATTG TTACGTGTAA TTTAAAATAT ATCTGTTTGC CATCGTCATG AGAAGATTAT 1380

ATGTAAGGCT CTGAAGGGAG AGGGAGATGT ACATTCTGCC AGGCTCCTGG GGACCTTATC 1440

CGAGTCATGA AATTGATGAC TGTTGATCCA GTGGTGCAAG AAGCTACACT CCATGTGTCA 1500

TCACGCTTAT GACTCCTAAT GTATTTTTAA GGCAAAAAAT GTCAGCCGAC TCCATCTTCA 1560

CCCCTCGATT CCTCGAGTCC AGCCTTTCTG TGCCAGTGCT TCACTGAGCC ACAACGCTCT 1620

CGCCATCGGG ACCCGGCTGG GCCTGGAGTC TCGGGGCACA GTTGCCATGG AGCCCTCCTG 1680

GGTCATTCTA CAAATGTGCT GAGTGCCAGC TGAAAACCCC ACAGGAGATG GAGTACCTTG 1740

GCCAAGCTTA AAGAGAAGAT TTTCTCAGGG TATTTATTAG TGTGTCCAGC AGGGTCAGGA 1800

AGCAGGATGG AAAGATGCAT TCAGACTGTT AATTTATTAA CAAGGCAAAT GATTTTGTGT 1860

TTCTTGATGA CAGACTATTA AGTTTGGGAC TTATTTTCCC ATTTGAGAAG TTATAATATA 1920

TATTTAAGAT GATAAGTTTC CTGCTTAAGT TGTGCCTTTC AGCTTCAATG AGTTTAAGGA 1980

GCACTAAGGG TAATGATACC AATGAGGGTT GGTTTATTAT CAAACCTGAA TAGCTGTGGT 2040

TTCTCCAGTA AATATTTTCT TCTACTGAAC ATGGAGCCAT TATTAAGAGT TGTGTGTTTT 2100

TTATTATGTA CATTTGTATA TTTTTTTGCT TGTTTGATGT TCTATTTTTC TAATAGTTTT 2160

CTTTTAGTTT CTTAAAGTTG TGATACTAGA TTTAGATTCT GATGCTAACT GCAAATCAGG 2220

TTGGTCTCTG CTGGGTCTCT CCTGCTTTTA TTTTACTTTA AGGACAAGTG TAGTTGTCGT 2280

CCACCACCTT TCAAAAAATG TGAAACTGCC CTGCCTCCCC TTTTTGCTGA CAACACTGTG 2340

TACATTGACC ACTTCCTACC ATACTTTATG TTGTAAAATC AAACTCTTTT GTGGTACATT 2400

ATCTCATGCT TCTGCAAATT CGAATAAATT CTATGGCTTC CAAAAAAAAA AAAAAAAAAA 2460

AAAAAAAAAA AAA 2473

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Val Ile Asp Ala Phe Ser Pro Gly Cys Leu Arg Lys Val Ala Lys Cys
1 5 10 15

Ile Leu Leu Ser Glu Leu Leu Arg Ser Pro Pro Thr Phe Gly Phe Cys
20 25 30

Ser Ser Lys Asn Ser Phe His Cys Asp Leu Phe Ser Leu Ser Gly Ser
35 40 45

Gln Pro Pro Gly Pro Leu Cys

50 55

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4093 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCTAATTTTT TGTATTTTTA GTAGAGATGG GGTTTCACCA CACTGGCCAG GCTTGTCTCG 60

CTCGAACTTC TGACCTCATG ATCCACCTGC CTCGGCCTCC CAAAGTGCTG GGATTACAGG 120

TGTGAGCCAC CGCACCTGGC AAAAGAGAAT CTTACAGAAC CTATTCACTG GGAAGGAAGC 180

CCTCATTATA ATGATTTTCA TTCTTATGTG TGTTTCAGGA CGACTGGGTT TGGATTCAGA 240

AGAGGATTAT TATACACCAC AAAAGGTGGA TGTTCCCAAG GCCTTGATTA TTGTTGCAGT 300

TCAATGTGGC TGTGATGGGA CATTTCTGTT GACCCAGTCA GGCAAAGTGC TGGCCTGTGG 360

ACTCAATGAA TTCAATAAGC TGGGTCTGAA TCAGTGCATG TCGGGAATTA TCAACCATGA 420

AGCATACCAT GAAGTTCCCT ACACAACGTC CTTTACCTTG GCCAAACAGT TGTCCTTTTA 480

TAAGATCCGT ACCATTGCCC CAGGCAAGAC TCACACAGCT GCTATTGATG AGCGAGGCCG 540

GCTGCTGACC TTTGGCTGCA ACAAGTGTGG GCAGCTGGGC GTTGGGAACT ACAAGAAGCG 600

TCTGGGAATC AACCTGTTGG GGGGACCCCT TGGTGGGAAG CAAGTGATCA GGGTCTCCTG 660

CGGTGATGAG TTTACCATTG CTGCCACTGA TGATAATCAC ATTTTTGCCT GGGGCAATGG 720

TGGTAATGGC CGCCTGGCAA TGACCCCCAC AGAGAGACCA CATGGCTCTG ATATCTGTAC 780

CTCATGGCCT CGGCCTATTT TTGGATCTCT GCATCATGTC CCGGACCTGT CTTGCCGTGG 840

ATGGCATACC ATTCTCATCG TTGAGAAAGT ATTGAATTCT AAGACCATCC GTTCCAATAG 900

CAGTGGCTTA TCCATTGGAA CTGTGTTTCA GAGCTCTAGC CCGGGAGGAG GCGGCGGGGG 960

CGGCGGTGGT GAAGAAGAGG ACAGTCAGCA GGAATCTGAA ACTCCTGACC CAAGTGGAGG 1020

CTTCCGAGGA ACAATGGAAG CAGACCGAGG AATGGAAGGT TTAATCAGTC CCACAGAGGC 1080

CATGGGGAAC AGTAATGGGG CCAGCAGCTC CTGTCCTGGC TGGCTTCGAA AGGAGCTGGA 1140

AAATGCAGAA TTTATCCCCA TGCCTGACAG CCCATCTCCT CTCAGTGCAG CGTTTTCAGA 1200

ATCTGAGAAA GATACCCTGC CCTATGAAGA GCTGCAAGGA CTCAAAGTGG CCTCTGAAGC 1260

TCCTTTGGAA CACAAACCCC AAGTAGAAGC CTCGTCACCT CGGCTGAATC CTGCAGTAAC 1320

CTGTGCTGGG AAGGGAACAC CACTGACTCC TCCTGCGTGT GCGTGCAGCT CTCTGCAGGT 1380

GGAGGTTGAG AGATTGCAGG GTCTGGTGTT AAAGTGTCTG GCTGAACAAC AGAAGCTACA 1440

GCAAGAAAAC CTCCAGATTT TTACCCAACT GCAGAAGTTG AACAAGAAAT TAGAAGGAGG 1500

GCAGCAGGTG GGGATGCATT CCAAAGGAAC TCAGACAGCA AAGGAAGAGA TGGAAATGGA 1560

TCCAAAGCCT GACTTCGATT CAGATTCCTG GTGCCTCCTG GGAACAAACT CCTGTAGACC 1620

CAGCCTCTAT TCTCCTGAGC CTATAGAGCC CCCAGGAGAC TGGGACCCAA AGAACTTCAC 1680

AGCACACTTA CCGAATGCAG AGAGCAGCTT TCCTGGCTTT GTTCACTTGC AGAAAAGGAG 1740

CGCAAGGCAG AGGCTCTGAA GCACTTTCCT TGTACATTTG GAGAGTGGCA TTGCCTTTTA 1800

GATAGGATCT AGGAGTGATT TTATTGTTTT GGAGAATGGA AGGGCCCCCA TGGCCCTGGC 1860

TTTGTCATCA GTGACTGCCA TAGCAACAGC AGCTCTGTAC CTCATCTGTT GATCCCACCT 1920

TTGAAGAGGA GACACAGTGC TCACCTTAAT TGCGCTGGTA GCAGCTTATA TCCCATGTAT 1980

CATTTTCACC ATTGATTGGA AGCTGCCTTG GGAATTCAGT ACCAGGCATT ACCCCTCTGG 2040

GTGGGAGAGG GAGAAGTGTA AAGTTGGAGT GGGCTGGAAT CAGGTGTGGC CCGCCCAGTG 2100

TCCTCTGCAG AGTGGTGAAG TAGTCTGGCC CTCTTGGGAG CCCTGAGTCC AGGAAAATAT 2160

GTCTGATGGA GTCAATCTAG GGCTTGTTTC GAAAAAGTTC AGTTACTCTG TGCAGCTAAA 2220

TGCTTTAGGA GGAAAGGTAG GCTTAGGTTG CTTTTCCTCT GAGGGTTGAT TGAAATTTCT 2280

TCAGTGAGGA ATAGAGAAAG GGCAGGACCC TCATCATCAC ACAGCTGGTG TTTCAGGCTG 2340

TGACCAATGC AGGGTGGGAT TTCCTAACTG TGGATGAGGG GATGAGGTGT CTCTGAGGGA 2400

TGAGGTGTCT CAGAGAATTG AGTCCATGGG GCAGTCAGAA TAGCCTTAAG AGAAAATCAT 2460

GAAGGAGAAG AGGTCCTCCT TTAGCTGCCT CTACTTGGTA TCTTAGAGAG GGCTTAGAGG 2520

GCTCTCAGTC TTCTGCCCAT GAAAAGACTT CTTTGAGCCT CTGCCTTCAT GGCTCTTAGG 2580

GTTCTGATCT TGATATCAGC AGCCCCAACC ACTTTCTTTC TGAATGTCTA GTCAGTATTT 2640

TTCCCCTTTT GGTGTTTTAT GAAGCCATGT GGTAACGAAT GAATCTGTAT CATTTTTCCT 2700

ACCTGAGTGG CCCAAAGCCA GCACCAAACC CTGGGAGTCC CTGAAGCCTA ACAGAACAGG 2760

TAGAACTTGC AAAAGGAATT TGGCTGAGAG CTCACTTCTA ATCCTGTACT CACTGTGTCT 2820

TTGTAGATAG AAACAAGCTA GCTTTACCAA AGAGAAATAG TGTCATAGAA GAACAAAACT 2880

TCATATAGAA AGTTCTAGGC AAGTATTTGA TGGTTTCCTT AAGGATGTGA GCTTTGTATT 2940

TCCACCTAGC CTTGTAAAAT GTTCCTGTGG TATTTTGTGT CACACATCCT ACCTTTGATG 3000

AGTCTCACAT CCCACCTTCC TATAAACAGC TAAATTAATT TTGTTTCATC TTCCCCAGAC 3060

CAAAATGTTT GATAATCTTA TCAACATTGT GGGAGGTCTT GTGCAATGGA AATTTTGCCA 3120

TTTCTCCAAA ACTGGTGGCA TAAAGGCTGA TGCTTGGGGA GAACCCCATT GCTCGGGACA 3180

GGCAACTCTG TTCAATGGGA TCTTCTTTGG TTTGATGTTC CCATTGTTTT CTCAGTTCTG 3240

GGAAGCCTAG TACATTAGTA CTAATGTAAT CACTGAAACC TTTTCTTGAA ATAAGGGAAG 3300

CAGCCAAACT TTGATTAAAG TTGCAAGTTC TGGGGACTTG CGGGGGTTGT CATAAACTGT 3360

AACAGTGGGT TTTGGTTCAG CATGTAAATG CAACTTTGAT TTTCTTGAGG ACCGATTGAC 3420

CTGTCATGTC CCTGTATCCT CATGCTCATC ATCTCAGCAG GCCTGAGAGG CTGGGTCAGT 3480

TTGGGTGTTC ATCATGAGGA TTGCTTCTGC CATGGAGCTG ATGGACGTGG GCAGGTTGCT 3540

GAGAAGGTGG GGTGAAAGTG AGTGCCGGGG GTGGGTGAGT GCCCTGGTCT TGTTCATAGG 3600

GGAGCCTTTC CCTAGCAGTG GAACGCTGTG GTCATTTTCT CTAGCATATT CCCTTGGGAA 3660

GTCTAGATTT GCTATTAATC TGGCTGAGAA TCTAAGTTCT GTGCCTTAGA GACAGTTTGC 3720

ACTTTCCCAT ATTGTGCCTG GGACAGCCAT ATGATTTTTT TTCCCACCAA ACAAGTATGC 3780

AAACAGAAAC CAGTTTCAAA GGGGGATGGA GTAAAAGATG AGGCAGTAGA AATGCCTTTG 3840

AATGGTTTTT CTGTAGCTAA TTCTCTTTAA ATTTTGTCCT GCTTTTTTTC TTTATGCAGT 3900

GCTAGGTGTT TTAAGTTTTC TAGTAGTATT GCTTTTGAGT TACAGTATAA CCTGAGTTAC 3960

TCCTCTGCTC TAACATTGTT GCAGAAGAGT AACTCAGGTT ATTGTTAGCC AGGTTGCTTG 4020

AAAGGTTGAG AGTGGAGTGG TTTGGCATTT CTGTTTTAAA TAAACATTTA AGCTCTTAAA 4080

AAAAAAAAAA AAA 4093

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 522 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ile Phe Ile Leu Met Cys Val Ser Gly Arg Leu Gly Leu Asp Ser
1 5 10 15

Glu Glu Asp Tyr Tyr Thr Pro Gln Lys Val Asp Val Pro Lys Ala Leu
20 25 30

Ile Ile Val Ala Val Gln Cys Gly Cys Asp Gly Thr Phe Leu Leu Thr
35 40 45

Gln Ser Gly Lys Val Leu Ala Cys Gly Leu Asn Glu Phe Asn Lys Leu
50 55 60

Gly Leu Asn Gln Cys Met Ser Gly Ile Ile Asn His Glu Ala Tyr His
65 70 75 80

Glu Val Pro Tyr Thr Thr Ser Phe Thr Leu Ala Lys Gln Leu Ser Phe
85 90 95

Tyr Lys Ile Arg Thr Ile Ala Pro Gly Lys Thr His Thr Ala Ala Ile
100 105 110

Asp Glu Arg Gly Arg Leu Leu Thr Phe Gly Cys Asn Lys Cys Gly Gln
115 120 125

Leu Gly Val Gly Asn Tyr Lys Lys Arg Leu Gly Ile Asn Leu Leu Gly
130 135 140

Gly Pro Leu Gly Gly Lys Gln Val Ile Arg Val Ser Cys Gly Asp Glu
145 150 155 160

Phe Thr Ile Ala Ala Thr Asp Asp Asn His Ile Phe Ala Trp Gly Asn
165 170 175

Gly Gly Asn Gly Arg Leu Ala Met Thr Pro Thr Glu Arg Pro His Gly
180 185 190

Ser Asp Ile Cys Thr Ser Trp Pro Arg Pro Ile Phe Gly Ser Leu His
195 200 205

His Val Pro Asp Leu Ser Cys Arg Gly Trp His Thr Ile Leu Ile Val
210 215 220

Glu Lys Val Leu Asn Ser Lys Thr Ile Arg Ser Asn Ser Ser Gly Leu
225 230 235 240

Ser Ile Gly Thr Val Phe Gln Ser Ser Ser Pro Gly Gly Gly Gly Gly
245 250 255

Gly Gly Gly Gly Glu Glu Glu Asp Ser Gln Gln Glu Ser Glu Thr Pro
260 265 270

Asp Pro Ser Gly Gly Phe Arg Gly Thr Met Glu Ala Asp Arg Gly Met

```
        275                 280                 285

Glu Gly Leu Ile Ser Pro Thr Glu Ala Met Gly Asn Ser Asn Gly Ala
    290                 295                 300

Ser Ser Ser Cys Pro Gly Trp Leu Arg Lys Glu Leu Glu Asn Ala Glu
305                 310                 315                 320

Phe Ile Pro Met Pro Asp Ser Pro Ser Pro Leu Ser Ala Ala Phe Ser
                325                 330                 335

Glu Ser Glu Lys Asp Thr Leu Pro Tyr Glu Glu Leu Gln Gly Leu Lys
            340                 345                 350

Val Ala Ser Glu Ala Pro Leu Glu His Lys Pro Gln Val Glu Ala Ser
        355                 360                 365

Ser Pro Arg Leu Asn Pro Ala Val Thr Cys Ala Gly Lys Gly Thr Pro
    370                 375                 380

Leu Thr Pro Pro Ala Cys Ala Cys Ser Ser Leu Gln Val Glu Val Glu
385                 390                 395                 400

Arg Leu Gln Gly Leu Val Leu Lys Cys Leu Ala Glu Gln Gln Lys Leu
                405                 410                 415

Gln Gln Glu Asn Leu Gln Ile Phe Thr Gln Leu Gln Lys Leu Asn Lys
            420                 425                 430

Lys Leu Glu Gly Gly Gln Gln Val Gly Met His Ser Lys Gly Thr Gln
        435                 440                 445

Thr Ala Lys Glu Glu Met Glu Met Asp Pro Lys Pro Asp Phe Asp Ser
    450                 455                 460

Asp Ser Trp Cys Leu Leu Gly Thr Asn Ser Cys Arg Pro Ser Leu Tyr
465                 470                 475                 480

Ser Pro Glu Pro Ile Glu Pro Pro Gly Asp Trp Asp Pro Lys Asn Phe
                485                 490                 495

Thr Ala His Leu Pro Asn Ala Glu Ser Ser Phe Pro Gly Phe Val His
            500                 505                 510

Leu Gln Lys Arg Ser Ala Arg Gln Arg Leu
        515                 520
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1601 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCCGACGAGG CCTTCGACCA TGAGGTCTCC GCCTTCTTCC CCGCCAACCT GGACTTCCTG 60

TGCCTGCAGG AGGTGTTTGA CAAGCGAGCA GCCACCAAAT TGAAAGAGCA GCTGCACGGC 120

TACTTCGAGT ACATCCTGTA CGACGTCGGG GTCTACGGCT GCCAGGGCTG CTGCAGCTTC 180

AAGTGTCTCA ACAGCGGCCT CCTCTTTGCC AGCCGCTACC CCATCATGGA CGTGGCCTAT 240

CACTGTTACC CCAACAAGTG TWACKACSAT GCCCTGGCCT CTAAGGGAGC TCTGTTTCTC 300

AAGGTGCAGG TGGGAAGCAC ACCTCAGGAM CAAARAATCG TCGGGTACAT CGCCTGCACA 360

CACCTGCATG CCCCGCAAGA GGACAGCGCC ATCCGGTGTG GGCAGCTGGA CCTGCTTCAG 420

GACTGGCTGG CTGATTTCCG AAAATCTACC TCCTCGTCCA GCGCAGCCAA CCCCGAGGAG 480

CTGGTGGCAT TTGACGTCGT CTGTGGAGAT TTCAACTTTG ATAACTGCTC CTCTGACGAC 540

AAGCTGGAGC AGCAACACTC CCTGTTCACC CACTACAGGG ACCCCTGCCG CCTGGGGCCT 600

GGTGAGGAGA AGCCGTGGGC CATCGGTACT CTGCTGGACA CGAACGGCCT GTACGATGAG 660

GATGTGTGCA CCCCCGACAA CCTGCAGAAG GTCCTGGAGA GTGAGGAGGG CCGCAGGGAG 720

TACCTGGCGT TTCCCACCAG CAAGAGCTCG GGCCAGAAGG GGCGGAAGGA GCTGCTGAAG 780

GGCAACGGCC GGCGCATCGA CTACATGCTG CATGCAGAGG AGGGGCTGTG CCCAGACTGG 840

AAGGCCGAGG TGGAAGAATT CAGTTTTATC ACCCAGCTGT CCGGCCTGAC GGACCACYTG 900

CCAGTAGCCA TGCGACTGAT GGTGTCTTCG GGGGAGGAGG AGGCATAGAC CGTCCGGAGC 960

AGCGGGGCYT CTGCCAGCCC TTGCAGCTGC AGCCCATCCC TGGGCCATGT CCCCTCCATC 1020

GAGTGCCCGG TGCTTGGGGG AGGAGGGCAG GGACAGGGAG GGAGCCACAG TCAGTGCCCG 1080

GGAACCTGGA AGCTGCGCTG CTCTGCGCCT CTGGGCCTCA CTGTGGSCAG AGGAGTCAGG 1140

CCCGCCCCAG GAGCCTCCAG CTGCCTAACC AGTGCCATTC TTTCACAACA CGATTTTCTA 1200

CAAATCTACA GCACAACCGA GTTTGTAACC CGTGGGTTAG TATGAGGACC GGGTTCGTGT 1260

ACTCTCTGTA TCTCCTCTTA AGCTTCGTCC AGGGTTCTTT ATTTTTGTCT GCTGCCAATG 1320

TCGTCTCGCA TGCCTGCACC CTCGCATGCA CGCTGCCCGC ATGCCACGTG CCACGCTGTA 1380

GCCACAGACC CCTTGCTCGG GCCTCACCCA AGGCCAAACT CCAAACACAA TCAGAACCAG 1440

CCAAAGAAGC ACTTCCTGGG CACGGCCACC AGCTCTCCCG CCTCCAGTGT GGGCCGGCTC 1500

CTGCAGGGTC CGAGGGCTGC ATCTCTACCA GCCAGCCCAG GGCTCTTCCC AGGGTCTCGC 1560

ATTCAAGGGC AATTACATTT TAAAAAAAAA AAAAAAAAAA A 1601

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Asp Val Ala Tyr His Cys Tyr Pro Asn Lys Cys Xaa Xaa Xaa Ala
1 5 10 15

Leu Ala Ser Lys Gly Ala Leu Phe Leu Lys Val Gln Val Gly Ser Thr
20 25 30

Pro Gln Xaa Gln Xaa Ile Val Gly Tyr Ile Ala Cys Thr His Leu His
35 40 45

Ala Pro Gln Glu Asp Ser Ala Ile Arg Cys Gly Gln Leu Asp Leu Leu
50 55 60

Gln Asp Trp Leu Ala Asp Phe Arg Lys Ser Thr Ser Ser Ser Ser Ala
65 70 75 80

Ala Asn Pro Glu Glu Leu Val Ala Phe Asp Val Val Cys Gly Asp Phe
85 90 95

Asn Phe Asp Asn Cys Ser Ser Asp Asp Lys Leu Glu Gln Gln His Ser
100 105 110

Leu Phe Thr His Tyr Arg Asp Pro Cys Arg Leu Gly Pro Gly Glu Glu
115 120 125

Lys Pro Trp Ala Ile Gly Thr Leu Leu Asp Thr Asn Gly Leu Tyr Asp
130 135 140

Glu Asp Val Cys Thr Pro Asp Asn Leu Gln Lys Val Leu Glu Ser Glu
145 150 155 160

Glu Gly Arg Arg Glu Tyr Leu Ala Phe Pro Thr Ser Lys Ser Ser Gly
165 170 175

Gln Lys Gly Arg Lys Glu Leu Leu Lys Gly Asn Gly Arg Arg Ile Asp

180 185 190

Tyr Met Leu His Ala Glu Glu Gly Leu Cys Pro Asp Trp Lys Ala Glu
195 200 205

Val Glu Glu Phe Ser Phe Ile Thr Gln Leu Ser Gly Leu Thr Asp His
210 215 220

Leu Pro Val Ala Met Arg Leu Met Val Ser Ser Gly Glu Glu Glu Ala
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2274 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGCGTTGCTG GTGCGAGTCC TTAGGAACCA GACTAGCATT TCTCAGTGGG TTCCAGTATG 60

CAGCCGATTG ATACCTGTGT CTCCTACCCA AGGACAGGGG GACAGGGCTC TGTCTCGCAC 120

TTCCCAGTGG CCCCAGATGA GCCAGTCCCA AGCATGTGGT GGATCAGAAC AGATTCCTGG 180

AATAGACATA CAGCTGAATA GGAAGTATCA CACCACACGT AAGCTTTCTA CTACCAAAGA 240

TTCCCCACAG CCTGTTGAGG AGAAGGTTGG TGCTTTCACA AAGATAATAS AAGCCATGGG 300

ATTCACGGGA CCTTTGAAAT ACAGTAAATG GAAGATTAAG ATTGCGGCCC TGCGCATGTW 360

TACTAGCTGT GTGGAGAAAA CTGACTTCGA GGAATTCTTT CTAAGGTGTC AGATGCCTGA 420

TACATTCAAT TCATGGTTTC TTATAACCCT ACTCCACGTC TGGATGTGTC TAGTCCGAAT 480

GAAGCAGGAA GGCCGGAGTG GGAAGTACAT GTGTCGTATC ATAGTTCATT TTATGTGGGA 540

GGATGTTCAG CAGCGCGGCA GAGTCATGGG GGTTAATCCC TATATCCTGA AGAAGAACAT 600

GATCCTCATG ACAAATCATT TCTATGCAGC GATCTTGGGA TATGATGAGG GGATCCTTTC 660

AGATGATCAT GGGCTGGCCG CTGCCCTCTG GAGAACCTTC TTCAACCGGA AATGTGAAGA 720

CCCTCGACAT CTTGAATTGC TGGTAGAGTA TGTGAGGAAA CAGATACAGT ACCTGGACTC 780

CATGAACGGG GAGGATCTGC TTCTGACAGG GGAGGTGAGC TGGCGCCCTC TAGTGGAGAA 840

GAATCCTCAG AGCATCCTGA AGCCCCATTC TCCGACTTAC AACGACGAGG GACTTTGATG 900

GGCTGGGCCC TCCGCACGGC CCGCCAGCTG GCTTCGAGGA ACCTCCAGGA GAGAAGTGCC 960

TGTTGGTCCA GGACCCTGCA GAAAGTGGCC TGAACTGACC TCTGAACAGC ATCTGTCAAA 1020

TACCTGGCCC CATTTGTGTT GAGTTTCCTC TTAGTGTGCC CAGGAGTCTG ATCTGCTGGG 1080

GTACAGGGCT GGGAGAACCC CTAGCTCTCC CGGGGTGTCC TCTCCCTTAG GGGAAGCCCC 1140

GAGTGAGAGT CCCCCAGCAC ACACTCCCCA ACCCCCTCCA GCAACTACAT GTGACTGATA 1200

GCTTTTCCCA AAGGCCAAGG AAGGGATGGT GTAGGTTCAA AAGGGAAACC CCCCAGGGCC 1260

TGCTGTGGCC TAGGAGCAGA TTGTAATGCT GCCGAGTCCG GTCGGTGACC ACGCGTTGTC 1320

CCTCGGCTTT CAGCCATGGG GTTGAGTTGG CCATTAAAAG AAACAGAGAC TTCTCTCTGC 1380

CATGGCCCTT CTTTATTCCA GGGACTTAGA AACTTGCCTG AGATGGTGGA CGCAGTAATG 1440

AGGGCACCGC GCAGCTCAGT TAGAGACGGA GAAAGGGAAG AGGCTGGGAT GGTCTCTGCT 1500

GCTCTTGCCT CTAGTTCATG GAGATGTGTC TCTGTTCAGG CCAAGATACA GCCAGCCAGG 1560

CCTGTCGTCT GGGACCCAGG AGGCCTCTGA TGACCAAGGG CTTTCACATC CTAAGTCATT 1620

TGGAAGGAGG CCTTGAGAAC AAAGTCACCT TTGTCACTCC CAGTGAACTG AATGAGGAAC 1680

ATGCTGTCTC CTGTCTTGGC CTCCCCTTTC ATGAGATACT GGGGAGAAGA GAACATTCCT 1740

CCTGGCTTAG TTGTAGCAGA CCCAGACCTG TGCCCAGCTT TGGTCCCCCT TCCCAACTTC 1800

TGAAGCACGT GCTGCAGAGC CACCTTGGTC TGAGCACCTG AGGACCAGCC CCTCCTCCCT 1860

CAGTGCGGGT CATCTCTTGG GGGATTTTCT TAAAGTGAAG AAAGGGGGTG GGGAACCATA 1920

TTGCCCCTCC CTCCCCCATC AAACTTCCTT CATTTAACTT GCTATAAAAT GAGTCATATA 1980

AAGAAACTCT ATATGGGTGA GGTATATCCC ACTTCTGTGA AAACATTACA AATCAAACCG 2040

CTTCTCTCAG TTTATTTAAG ATGCTTTTGT TGCGAGCGGA GCTCTAGAGT GAAGCCTCCT 2100

GTGTGTGTGT GAGATAATAA CACCTTGTAA CTCATTACAG CTGGGCACTA TTTACATAAA 2160

CCAGAGCTGA GCCAGGCAGG AATTTGCTGA TTAATTTATT TTTAATGGAG TGAAGTATAC 2220

CATGCACCAA AATAAACTTT ACTGTGTGTA CCTAAAAAAA AAAAAAAAAA AAAA 2274

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 253 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Ser Gln Ser Gln Ala Cys Gly Gly Ser Glu Gln Ile Pro Gly Ile
1               5                  10                  15

Asp Ile Gln Leu Asn Arg Lys Tyr His Thr Thr Arg Lys Leu Ser Thr
            20                  25                  30

Thr Lys Asp Ser Pro Gln Pro Val Glu Glu Lys Val Gly Ala Phe Thr
        35                  40                  45

Lys Ile Ile Xaa Ala Met Gly Phe Thr Gly Pro Leu Lys Tyr Ser Lys
    50                  55                  60

Trp Lys Ile Lys Ile Ala Ala Leu Arg Met Xaa Thr Ser Cys Val Glu
65                  70                  75                  80

Lys Thr Asp Phe Glu Glu Phe Phe Leu Arg Cys Gln Met Pro Asp Thr
                85                  90                  95

Phe Asn Ser Trp Phe Leu Ile Thr Leu Leu His Val Trp Met Cys Leu
            100                 105                 110

Val Arg Met Lys Gln Glu Gly Arg Ser Gly Lys Tyr Met Cys Arg Ile
        115                 120                 125

Ile Val His Phe Met Trp Glu Asp Val Gln Gln Arg Gly Arg Val Met
    130                 135                 140

Gly Val Asn Pro Tyr Ile Leu Lys Lys Asn Met Ile Leu Met Thr Asn
145                 150                 155                 160

His Phe Tyr Ala Ala Ile Leu Gly Tyr Asp Glu Gly Ile Leu Ser Asp
                165                 170                 175

Asp His Gly Leu Ala Ala Ala Leu Trp Arg Thr Phe Phe Asn Arg Lys
            180                 185                 190

Cys Glu Asp Pro Arg His Leu Glu Leu Leu Val Glu Tyr Val Arg Lys
        195                 200                 205

Gln Ile Gln Tyr Leu Asp Ser Met Asn Gly Glu Asp Leu Leu Leu Thr
    210                 215                 220

Gly Glu Val Ser Trp Arg Pro Leu Val Glu Lys Asn Pro Gln Ser Ile
225                 230                 235                 240

Leu Lys Pro His Ser Pro Thr Tyr Asn Asp Glu Gly Leu
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TATCCATTAC GTCGACTAAT ACGTACATAA GAATTCAATC GGGCCTTGGG GCTGGCCCTG 60

AAACCTGCGA GGGGCTTCCG TCCACGTCCC CAGTGGACCT ACCACCCCTC CATCTGGGAA 120

AGCAGGCCAC AGCAGCCGGA CAAAGGAAGC TCCTCAGCCT CTAGTCGCCT CTCTGTGCAT 180

GCACATCGGT CACTGATCTC GCCTACTGGC ACAGACGTGT TTATCGGCCA ACCTGACCCT 240

CACAAAAAGC TACCACCGAA GTGGACAGGC CCCTACACTG TGATACTCAG CACACCAACT 300

GCAGTGAGAG TCCGAGGACT CCCCAACTGG ATCCATCGCA CCAGGGTCAA GCTCACCCCC 360

AAGGCAGCTT CTTCCTCCAA AACATTAACA GCTAAGTGTT TGTCTGGGCC AATTTCTCCT 420

ACCAAGTTTA AATTAACCAA CATTTTTTTC TTAAAACCAA AACACAAGGA AGACTAACCA 480

CGTGCTTCCA GGAATGGCCT GTATCTACCC AACCACTTTC TATACCTCTC TTCCAACCAA 540

AAGTCTTAAT ATGGGAATAT CCCTCACCAC GATCCTAATA CTGTCAGTAG CTGTCCTGCT 600

GTCCACAGCA GCCCCTCCGA GCTGCCGTGA GTGTTATCAG TCTTTGCACT ACAGAGGGGA 660

GATGCAACAA TACTTTACTT ACCATACTCA TATAGAAAGA TCCTGTTATG GAAACTTAAT 720

CGAGGAATGT GTTGAATCAG GAAAGAGTTA TTATAAAGTA AAGAATCTAG GAGTATGTGG 780

CAGTCGTAAT GGGGCTATTT GCCCCAGAGG GAAGCAGTGG CTTTGCTTCA CCAAAATTGG 840

ACAATGGGGA GTAAACACTC AGGTGCTTGA GGACATAAAG AGAGAACAGA TTATAGCCAA 900

AGCCAAAGCC TCAAAACCAA CAACTCCCCC TGAAAATCGC CCGCGGCATT TCCATTCCTT 960

TATACAAAAA CTATAAGCAG ATGCATCCCT TCCTAAGCCA GGAAAAAATC TGTTTGTAGA 1020

TCTAGGAGAA CCATTGTGCT TACCATGAAT GTGTCCAATT GTTGGGTATG CGGGGGAGCT 1080

TTATGAGTGA ACAGTGGCTG TGGGACGGGA TAGACATTCC CCCTTACTTA CAGGCATCCC 1140

AAAACCCCAG ACTCACTTTC ACTCCTCAGG AATGCCCGCA GTCCTGGACA CTTACCAACT 1200

CAGTATGAGG GACGGTGTGC ATATCCCGCA AGTGGACTGA TAAAACCCAT CGCGCCGTAG 1260

GTGAAAACCC GTCACCAAAC CCTAACAGTC AATGCCTCCA TAGCTGAGTG GTGGCCAAGG 1320

TTACCCCCTG GAGCCTGGTC TCCTTCTAAC TTAAGCTACC TCAATTGTGT CTTGTCAAAA 1380

AAGGCCTGGT ACTGTACAAA CACCACTAAC CCTTATGCCG CATACCTCCG CCTAAGTGTA 1440

CTATGCGACA ATCCTAGGAA CACCAGCTGA CAATGGACTG CCACTGACGG ATTCCTGTGG 1500

ATATGGGGAA CCCAGGCTTA CTCACAGCTA CCTTATCACT GGCAAGGTAC TTGCTTCCTA 1560

GGCACAATTC AACCTGGATT CTTTTTACTT CCGAAGCACG CGGGCAACAC CCTCAGAGTC 1620

CCTGTGTATG ATAACCAGAG AAAAATGATC CTTGGAGGTA GGAGGGAGCC AAAGATTGTG 1680

AGAGGACGAG TGGCCTCCGC AACGGATCAT TGAATACTAT GGTCCTGCCA CTTGGGCAGA 1740

GGATGGTTCA TGGGGTTATC GCACTCCCAT ATATATGCCA AATAGAGCGA TTAGACTACA 1800

AGCTGTTCTA GAGATAATCA CTAACCAAAC TGCCTCAGCC CTAGAAATGC TCGCGCAACA 1860

ACAAAACCAA ATGCGCGCGG CAATTTATCA AAACAGGCTG GCCCTAGACT ACTTATTAGC 1920

AGAAGAGGGT GCGGGCTGTG GTAAGTTTAA CATCTCCAAT TGCTGTCTTA ACATAGGCAA 1980

TAATGGAGAA GAGGTTCTGG AAATCGCTTC AAACATCAGA AAAGTAGCCC GTGTACCAGT 2040

CCAAACCTGG GAGGGATGGG ACCCAGCAAA CTTCTAGGAG GGTGGTTCTC TAATTTAGGA 2100

GGATTTAAAA TGCTGGTGGG GACAGTCATT TTCATCACTG GGGTCCTCCT GTTTCTCCCC 2160

TGTGGTATCC CATTAAAACT CTTGTTGAAA CTACAGTTAA CCTCCTGACA ATCCAGATGA 2220

TGCTCCTGCT ACAGCGGCAC GATGGATACC AACCCGTCTC TCAAGAATAC CCCAAAAATT 2280

AAGTTTTTCT TTTTCCAAGG TGCCCACGCC ACCCYTATGT CACGCCTGAA GTAGTTATTG 2340

AGAAAGTCGT CCCTTTCCCC TTTTCTATAA CCAAATAGAC AGGAATGGAA GATTCTCCTC 2400

GGGGCCTGAA AGCTTGCGGG ATGAATAACT CCTCCTCCTC AGGCCCAGTC CCAAGGTACA 2460

AACTTGCACC AGCAGCAAGA TAGCAGAGGC AGGAAGAGAG CTGGCTGGAA GACACGTACC 2520

CCCTGAAGAT CAAGAGGGAG GTCGCCCTGG TACTACATAG CAGTCACGTT AGGCTGGGAC 2580

AATTCCTGTT TACAGAGGAC TATAAAACCC CTGCCCCATC CTCACTTGGG GCTGATGCCA 2640

TTTTAGGCCT CAGCCTGTCT GCATGCAGGC GCTCATTAAA ACAGCATGTT GCTCCAAAAA 2700

AAAAAAAAAA A 2711

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 160 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Ala Cys Ile Tyr Pro Thr Thr Phe Tyr Thr Ser Leu Pro Thr Lys
1               5                  10                  15

Ser Leu Asn Met Gly Ile Ser Leu Thr Thr Ile Leu Ile Leu Ser Val
            20                  25                  30

Ala Val Leu Leu Ser Thr Ala Ala Pro Pro Ser Cys Arg Glu Cys Tyr
        35                  40                  45

Gln Ser Leu His Tyr Arg Gly Glu Met Gln Gln Tyr Phe Thr Tyr His
    50                  55                  60

Thr His Ile Glu Arg Ser Cys Tyr Gly Asn Leu Ile Glu Glu Cys Val
65                  70                  75                  80

Glu Ser Gly Lys Ser Tyr Tyr Lys Val Lys Asn Leu Gly Val Cys Gly
                85                  90                  95

Ser Arg Asn Gly Ala Ile Cys Pro Arg Gly Lys Gln Trp Leu Cys Phe
            100                 105                 110

Thr Lys Ile Gly Gln Trp Gly Val Asn Thr Gln Val Leu Glu Asp Ile
        115                 120                 125

Lys Arg Glu Gln Ile Ile Ala Lys Ala Lys Ala Ser Lys Pro Thr Thr
    130                 135                 140

Pro Pro Glu Asn Arg Pro Arg His Phe His Ser Phe Ile Gln Lys Leu
145                 150                 155                 160
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2892 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGATTATTGA GCATAACTGA TGAGAGAATG ACAGTATTGG TGACTAGCTT TTTTGTTTTT 60

AAATAACACA CTTTAAAATG CCACAGTGTT ATTTGAAAGC TTAAAGGCAT TTTCTCATCT 120

TCAGTATTTG CTTTTTATCA GTTGTAGAAA AGAACCTTAG TGTTTTCTTC TCTACCTATG 180

TATGCTTATG AAACCGCTCT ACCTACATAT GAAACTTCTC TACCTATGTA TGTTTACGAA 240

AAGAGATGTA TTTGCCAAGA AAATCTTGAT TATATAAAAG ACAAAAAGAT TATAAAAGAC 300

AGTTTCTGTT TAAGTAAAAC TGTCTTTGAA ACTTAAGAAA GCTTAAGTTT TAAAAAAATC 360

TCAGTTAAAC ATCATTGGTT TATTTGACTT ACTGTTAACT CTTGCTTCTC TTTGACTCCA 420

TGGTGTTTTT AGATAACCAT GTATGAACTT TATGTTAATG GTTTGTGGGA TACTAAAATA 480

GCTTGAAGAT GACTTGATGA CTGTGCATTT TATATAGTTT TATTCTCCTA AAATCTCAGG 540

AGGGCAGCAA GTGCTGTCAA CGATTATATA GTGATGAGAT TTTTATGGGA ATGATTTCTT 600

CTTGGTGCGT TTTACACATT TGTACTGATA GCAAAACTAA AGTTTAAAGC AGCAAAGTTT 660

AAGTTCTCCT TAAATCCTAC AGAAACCAAC CTTTTAAGGA CATAATTTCA TCTAAAACAT 720

GACGATATTT AGCACACCTT TTAATGTGGG TATATATCAA GTGTTTAAGG ACTGGCTAGT 780

ATGTGATAGA GCAAGACCTG AGACTTTATA AGTATTTGCT CGTGTCTGTT GACAGACCTC 840

TTTCTTTCAA ACTTGTTAGA AGAGTGGTAA GACATATCCA ATTGGAAAAT AAGATGCAGT 900

GTTGTATAGC ACATACATTT AAAGTGCTTG CGTTAAAATT AGTTTCTCAA TAAGATAAAA 960

TTATTTTAAA AATTTGGTTC ACTTTATTAC AATAGTGGCA ATTTAGCTTT TCAGTATTAC 1020

AGGAATTTAA AAATTGGTTT CTTGTAGGGG ACATCTCAAC TTTGGGAATA TCTTCACTTA 1080

ATTTTTAAAA AATATTTTCA TGCTTTATTG TCCAGCTATA CAATATATCG CAAAATCCTG 1140

ACAAGTTCAT TGTATTAAGG TATTAACTAT TACATGGAAA GCAATTCTGT TCATCTTTTG 1200

ATGTTTGTGT TGAAAATGCT TATCTTTGTG TTTTGATCTT CCAACAGCTG AGAGCTTGAA 1260

CTGATTTAAA CATTTGTCAA TATACTTAAG AATGCTTTAA GTAAAGAAGG GGAAAATTTT 1320

AAGTAAGTTT TTTCCCTTCT AGGAAGAAAA ACTATGATGA TGTTAAGAAA ATGTCATTAT 1380

AGAGCTTGCT CAATAATATG TTCTTTAAAT CCACCTCCAT TTGTACATTA TAGGTATCAT 1440

TCTGTTTTTG CTTAAAATAA TCTGCAACCA TTTCAGATAG TTTTACAGCA AATTGATCTA 1500

AAAGCCACTA ATAAATTCTA GGGTTTGAGT CTAGAAGCCA AGCAAACTGT CACCAATGTC 1560

AGTTGTAAAT TAGAATGCAA CATGAGGCTT CAGACTCATG ACAATGATAT ACATGAAAAC 1620

AAAAATATAA TTGTGTCTAC CTTCCTACTT TCCCTTTTGA CATATGTAGT TGGAATTTTA 1680

CATAGTCTTA AAATCCATAT TTAGAATCTT ACCTGTTTCT ATAATAATTA GTAAAATGCC 1740

AAAGTAGTGA TAGAATATTG TGGCATTGAA GTAGCCGAAA AATTGTTAGT TTTAGCATCA 1800

AAAAAGTAAA TAGATGTTGA AATGAATTTT TGTATGTGCC AGGTTGAAGA GAGTGTGCCA 1860

GTGACAGGAA GTAGTCTAAA AAATTAACAG TTATGGTTTT AATAGGATCT GAAAGACAAT 1920

CTTTAAAGAA ATGGGAGAAA TTGGGGGTAT CAGTGAACCT ATACCAACCT CTCTTTGTAC 1980

ATAAATATGG TGATGTAGCT AGATATAAAA ATCAGTGTCT TACTGGCACC ATTTACAGTT 2040

TAGAAAACAA TCTTTTTCTT AAAAATGCCC ATCTGATTTC TATTTTTAGG AGCTACTTGG 2100

ATTTGTATGT ATTTTTTCTA CGTGAAAATA TATGTACTCT TCACTTTTGT TCCAGTACTA 2160

TAATTGCTCA TGCACTCTTT CTCCCCTTTG AGAACATTCA GTGAAATACA ACTTCATCAA 2220

AGATTTGCTC AAAGGAGAAG AATCGCATGA GTGTGAAAAG TAGATGCTCG TAGCCAGAAC 2280

AGAAAAGGTT ACACATGATC ATGGCACAGA AGATAGGAGG TTTGACTTGG TGGGCCATAA 2340

TGTTTATTAT CCTTTTTGAA ATAACAGGGA CCAGCAGCAG TTTTCTCAGG ATAAATGCTC 2400

TACCCCACTT CTCTATGAAC AGGTGTGGGG AGGCTTACTT TCCATTTTCA TATTTATACA 2460

CCTCTCTACA AAAGCAATTT TTAATGAAGG TTAGTGGAAT TGTTAAAAAT CTGAGAGGAA 2520

TGATGACTGG AGGTGTTTGG GGTTTTTTTC TGTATTCATT TTTTAATGAG AAAAGTTTTA 2580

AATGTAGTAC AGGTTAGACC CAACTACTAC CTTACTATTA TAGGACGATT CTATGTTTCT 2640

GTTAAAGTAT TCAAGTAGCT TTCTCTGGGG GAAAAAGTAC CACTTGGACA CTTAAAGGAA 2700

TTGGGATTTT TGTCTACTTT GGATAAGGCA GTTGACTTCT TAAGTAAAAG CAATAGTGTA 2760

AAATGTCATT TTGTTTGGAA TGTTAAGTGA GCAAATAAAA AACATGTTGA AATTGTAAAA 2820

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2880

AAAAAAAAAA AA 2892

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 100 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ile Met Ala Gln Lys Ile Gly Gly Leu Thr Trp Trp Ala Ile Met
1 5 10 15

Phe Ile Ile Leu Phe Glu Ile Thr Gly Thr Ser Ser Ser Phe Leu Arg
20 25 30

Ile Asn Ala Leu Pro His Phe Ser Met Asn Arg Cys Gly Glu Ala Tyr
35 40 45

Phe Pro Phe Ser Tyr Leu Tyr Thr Ser Leu Gln Lys Gln Phe Leu Met
50 55 60

Lys Val Ser Gly Ile Val Lys Asn Leu Arg Gly Met Met Thr Gly Gly
65 70 75 80

Val Trp Gly Phe Phe Leu Tyr Ser Phe Phe Asn Glu Lys Ser Phe Lys
85 90 95

Cys Ser Thr Gly
100

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 618 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATTCGTTACA AACATGTTTG CATGCCTACC ATATCTCAGG CACTGGGGAT ACAGCAGATC 60

AAGATCCTAC CCCATGGAAC TAAAAGAGGA CAGAGTACTG AGTGGAACAT ACGATGATAG 120

ATTTACAAAT AATGTAGCAT ACTTCTACTT CATTGTATCT TAAGTTTCTT GAAATATTGC 180

TACTGGAGAT TGGAAAGAAA TCTTAATGTT ATGGGGTATT GTCTAAGAAG CTTTATTTTA 240

AAACCATCTC ATTAAATTTT GTTGCATTTT AGATAATCGT CCCCAGATGC CATGTTACCC 300

TAGTGCAGAG TTTGGGGCTG GATAAGTTTT TGTTGTAGGT GGCTATCCTG TGTTTTGTAG 360

GGTATTTAGC AGCATCCTGG CCTTAAAACA AAAATGTTTT CAGACATTGC CAAATGTCCC 420

CCGAGCGGTA AAGTCACCCC CAAGTTGAGA ACCGCTCTAT ACAAAGAGCT GTTATTAGAG 480

CTAGACATTT CTGAATTGGC ATCAATTTCT ATATTGTATC CATAAACATT AGTAGCCACG 540

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 600

AAAAAAAAAA AAAAAAAA 618

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Phe Ser Asp Ile Ala Lys Cys Pro Pro Ser Gly Lys Val Thr Pro
1 5 10 15

Lys Leu Arg Thr Ala Leu Tyr Lys Glu Leu Leu Leu Glu Leu Asp Ile
20 25 30

Ser Glu Leu Ala Ser Ile Ser Ile Leu Tyr Pro
35 40

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 772 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TGCCATACAC TTCAGCAGAG TTTGCAACTT CTCTTCTAAG TCTTTATCCT TCCCCCAAGG 60

CATGCCTAGC ACAGGACTCT TGAACAGTGA TGCCTCAATT AGAGTTGCTA GCCAATAGAT 120

TGAAGCTATG TTGGCACAAT ATCCTACATC CTCCCGATCT ACTGGCTGAG CCCAACCCCA 180

CCTAAGAAGG ACAATAAAGA TCTGTGTTCA GAGTCATACT GAATAGAGAC TTCTGGACTC 240

TATAGAACCC ACTGCCTCCT GATGAAGTCC CTACTGTTCA CCCTTGCAGT TTTTATGCTC 300

CTGGCCCAAT TGGTCTCAGG TAATTGGTAT GTGAAAAAGT GTCTAAACGA CGTTGGAATT 360

TGCAAGAAGA AGTGCAAACC TGAAGAGATG CATGTAAAGA ATGGTTGGGC AATGTGCGGC 420

AAACAAAGGG ACTGCTGTGT TCCAGCTGAC AGACGTGCTA ATTATCCTGT TTTCTGTGTC 480

CAGACAAAGA CTACAAGAAT TTCAACAGTA ACAGCAACAA CAGCAACAAC AACTTTGATG 540

ATGACTACTG CTTCGATGTC TTCGATGGCT CCTACCCGTT TCTCCCACTG GTTGAACATT 600

CCAGCCTCTG TCTCCTGCTC TAGGATCCCC GACTCATTAA AGCAAAGAGG CTTAAAAAAA 660

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 720

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 772

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 131 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Lys Ser Leu Leu Phe Thr Leu Ala Val Phe Met Leu Leu Ala Gln
1 5 10 15

Leu Val Ser Gly Asn Trp Tyr Val Lys Lys Cys Leu Asn Asp Val Gly
20 25 30

Ile Cys Lys Lys Lys Cys Lys Pro Glu Glu Met His Val Lys Asn Gly
35 40 45

Trp Ala Met Cys Gly Lys Gln Arg Asp Cys Cys Val Pro Ala Asp Arg
50 55 60

Arg Ala Asn Tyr Pro Val Phe Cys Val Gln Thr Lys Thr Thr Arg Ile
65 70 75 80

Ser Thr Val Thr Ala Thr Thr Ala Thr Thr Thr Leu Met Met Thr Thr
85 90 95

Ala Ser Met Ser Ser Met Ala Pro Thr Arg Phe Ser His Trp Leu Asn
100 105 110

Ile Pro Ala Ser Val Ser Cys Ser Arg Ile Pro Asp Ser Leu Lys Gln
115 120 125

Arg Gly Leu
130

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 875 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCTGTTTTCA ATATTTTATA ATGAAAATTT TTGAACATTC GGAGAAGTTG AAAGAATTAC 60

ACCCAGAACA CCCATGTCTA CCATGTTACT ACATTTGTGT GTGTGTGTGT GTGTGTGTGT 120

GTGTGCGTGT GCAATTTTAA TTACAATGGT CGAGGAAGCT CTTGCTGAGA AGGTGATGTT 180

GAATAAAGAC TCTAAAGACC CTAAAGAGTT GAGAGAGAGC TGTGTGGAGT TCTGGGGCCC 240

AGGCACAGCA AGTACAAAGA TCCTGAAGCA GGAGCATTCT TGGTGTGTTC AAGGAAAGCA 300

AGGAGGCCAG TGAGGTTGGA AAAGGGAATG AGGTCAGART AATAATAGGG TGAAAGARGA 360

TGGCTGGGGG ATGGGGGGGC ARTGARGCAG GGCCTATGGT TTCTACTTGG TGARGTGGGA 420

ARCCACTGGA RGGGTTTAAG CCGATAATTG ATGTCACATA ATTTATGTTG TAATGGAACC 480

CGTGTGACTG CTCCTGGGGA ACAGACAAAA GAAAGGGTAG TAGAGACACC AGCTAGGAAG 540

CAGATTCAGC TAGAAAGAT GATACCTTCC ACTAAGGTGT TGGAGAAGTG GTTGGATTCT 600

AGATAATTTT TGAAGGTGGA GTTGGCAGAA TTTGAAGATC ATTCCATTTC TGTTCATACA 660

GAGCTTCCTC ATTCTCTTTT GACAGCTGCC TAGAATCTCA TTGTATCATA ATGTTTTAAA 720

CTAGTCCCCG ATGATGAATA TTTAGGTTGA TGCTTTCTCT CTTGCTGCCA CAAAACAGAC 780

ATATATTTGT ATAAAATGCC ACGTGGACAT GTCATTCTTC ACATAAGTGA GTATTTGTAG 840

AATAAATTCC AAGAAGCAGA AAAAAAAAAA AAAAA 875

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 79 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Ser Thr Met Leu Leu His Leu Cys Val Cys Val Cys Val Cys Val
1 5 10 15

Cys Ala Cys Ala Ile Leu Ile Thr Met Val Glu Glu Ala Leu Ala Glu
20 25 30

Lys Val Met Leu Asn Lys Asp Ser Lys Asp Pro Lys Glu Leu Arg Glu
35 40 45

Ser Cys Val Glu Phe Trp Gly Pro Gly Thr Ala Ser Thr Lys Ile Leu
50 55 60

Lys Gln Glu His Ser Trp Cys Val Gln Gly Lys Gln Gly Gly Gln
65 70 75

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TNCACGATGCC CAGTGCAAGC AGGACCAC 29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TNGGACAAGGG TGTGTGAGCA GGGATGAT 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GNGCTTCCTTC CCAGTGAATA GGTTCTGT 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ANAAACTCGGT TGTGCTGTAG ATTTGTAG 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ANGAACTAGAG GCAAGAGCAG CAGAGACC 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ANCTCTTTCCT GATTCAACAC ATTCCTCG 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CNCTCATGCGA TTCTTCTCCT TTGAGCAA 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ANCTCTGCACT AGGGTAACAT GGCATCTG 29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNGTGAACAGT AGGGACTTCA TCAGGAGG 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ANTGTAGTAAC ATGGTAGACA TGGGTGTT 29

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 65 to nucleotide 1270;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1139 to nucleotide 1270;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1011 to nucleotide 1216;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BO114_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BO114_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BO114_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BO114_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 196 to amino acid 205 of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 326 to amino acid 384;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 196 to amino acid 205 of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone BO114_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

9. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

10. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 326 to amino acid 384.

11. A composition comprising the protein of claim 8 and a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 418 to nucleotide 582;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 508 to nucleotide 582;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 555;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CD311_2 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CD311_2 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CD311_2 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CD311_2 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 22 to amino acid 31 of SEQ ID NO:4;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

15. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 46;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 22 to amino acid 31 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone CD311_2 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

17. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 191 to nucleotide 1756;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 254 to nucleotide 1756;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 604;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CG279_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CG279_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CG279_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CG279_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 256 to amino acid 265 of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

18. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 138;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 256 to amino acid 265 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone CG279_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

20. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 226 to nucleotide 948;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1128 to nucleotide 1601;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CJ424_9 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CJ424_9 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CJ424_9 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CJ424_9 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 115 to amino acid 124 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

21. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 115 to amino acid 124 of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone CJ424_9 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

23. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 137 to nucleotide 895;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1488 to nucleotide 2274;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CR930_1 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CR930_1 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CR930_1 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CR930_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 121 to amino acid 130 of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 121 to amino acid 130 of SEQ ID NO:10; and

(c) the amino acid sequence encoded by the cDNA insert of clone CR930_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

26. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 494 to nucleotide 973;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 611 to nucleotide 973;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 521 to nucleotide 940;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DA306_4 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DA306_4 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DA306_4 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DA306_4 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 75 to amino acid 84 of SEQ ID NO:12;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

27. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 11 to amino acid 149;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 75 to amino acid 84 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone DA306_4 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.

29. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 2295 to nucleotide 2594;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1867 to nucleotide 2372;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DG76_1 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DG76_1 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DG76_1 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DG76_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 45 to amino acid 54 of SEQ ID NO:14;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

30. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 26;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 45 to amino acid 54 of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone DG76_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

32. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 394 to nucleotide 522;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 476;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone DO19_1 deposited under accession number ATCC 98333;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone DO19_1 deposited under accession number ATCC 98333;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone DO19_1 deposited under accession number ATCC 98333;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone DO19_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

33. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 27;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone DO19_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

35. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 262 to nucleotide 654;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 322 to nucleotide 654;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 618;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone EQ219_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone EQ219_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone EQ219_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone EQ219_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 60 to amino acid 69 of SEQ ID NO:18;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

36. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 119;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 60 to amino acid 69 of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone EQ219_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

38. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 74 to nucleotide 310;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 125 to nucleotide 310;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 1 to nucleotide 338;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone FG340_1 deposited under accession number ATCC 98333;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone FG340_1 deposited under accession number ATCC 98333;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone FG340_1 deposited under accession number ATCC 98333;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone FG340_1 deposited under accession number ATCC 98333;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 34 to amino acid 43 of SEQ ID NO:20;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

39. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 75;

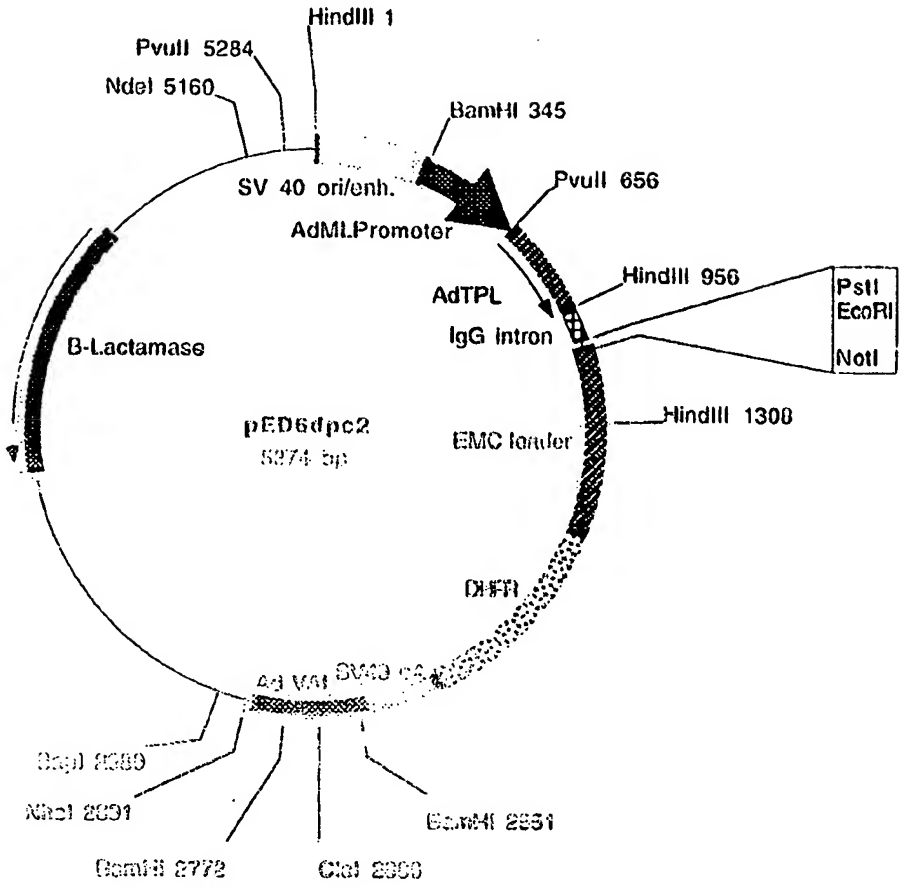
(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising the amino acid sequence from amino acid 34 to amino acid 43 of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone FG340_1 deposited under accession number ATCC 98333;

the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

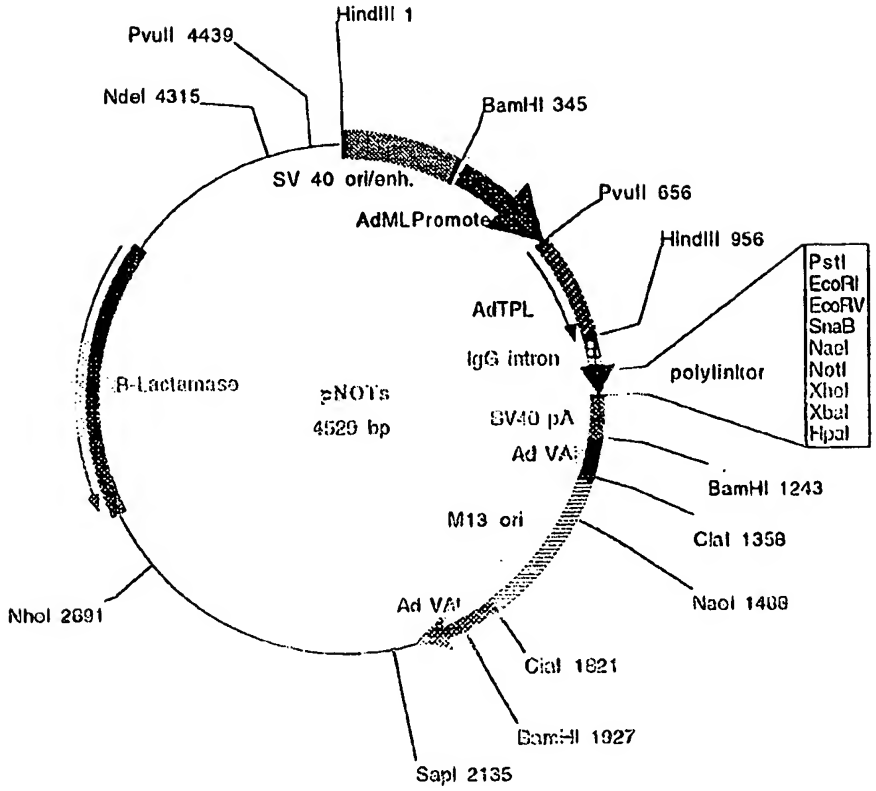
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI